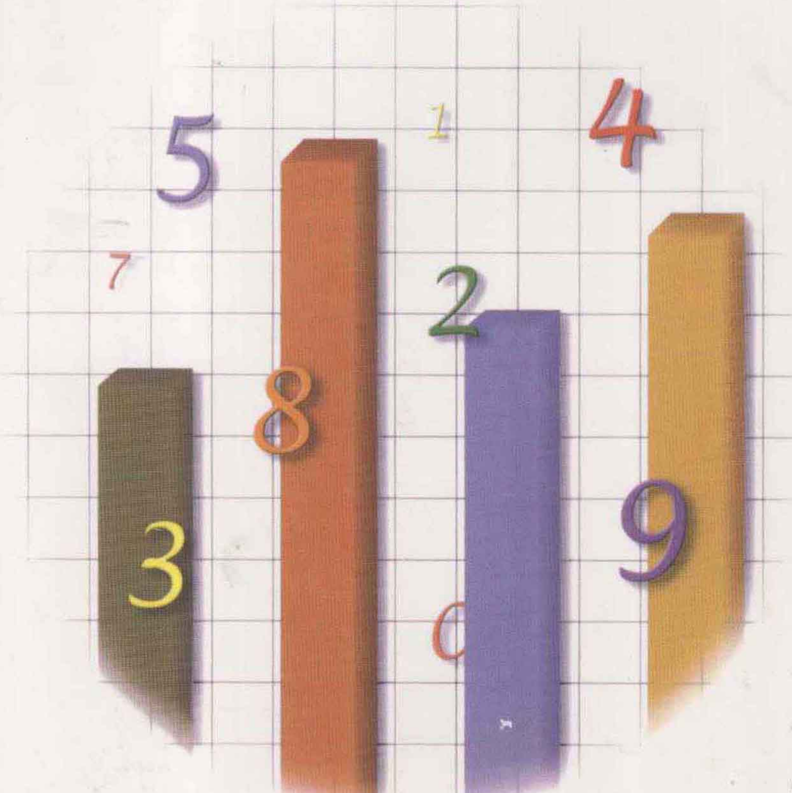
NetEffect Series

Technology Skills

A QUICK COURSE IN STATISTICAL PROCESS CONTROL

MICK NORTON



- TRAIN
- COMMUNICATE
- PERFORMANCE

A Quick Course in Statistical Process Control

MICK NORTON
College of Charleston

NetEffect Series

Upper Saddle River, New Jersey
Columbus, Ohio

Library of Congress Cataloging-in-Publication Data

Norton, Mick.
A quick course in statistical process control / Mick Norton.
p. cm.
Includes bibliographical references and index.
ISBN 0-13-093062-8
1. Quality control—Statistical methods. I. Title.
TS156.N65 2005
658.5'62—dc22

2004005801

Executive Editor: Debbie Yarnell
Production Editor: Louise N. Sette
Production Supervision: Carlisle Publishers' Services
Design Coordinator: Diane Ernsberger
Cover Designer: Kristi Holmes
Production Manager: Deidra Schwartz
Marketing Manager: Jimmy Stephens

This book was set in Goudy WtcTReg by Carlisle Communications, Ltd. It was printed and bound by R. R. Donnelley & Sons Company. The cover was printed by The Lehigh Press, Inc.

Pearson Education Ltd.
Pearson Education Singapore Pte. Ltd.
Pearson Education Canada, Ltd.
Pearson Education—Japan
Pearson Education Australia Pty. Limited
Pearson Education North Asia Ltd.
Pearson Educación de Mexico, S.A. de. C.V.
Pearson Education Malaysia Pte. Ltd.

[redacted]

A Quick Course in Statistical Process Control

Preface

This book is aimed at those who come in contact in any way with statistical process control (SPC) or people who simply want an introduction to the subject. The first four chapters examine why control charts work, how to update charts, how to interpret what the charts are saying, how to set up a chart, and how to assess the capability of a process being charted. Besides being of interest to those engaged in any of these specific activities, these chapters also should give decision makers some ideas about what kinds of variables they can track or control. Just enough probability is introduced, lightly, for the reader to understand how infrequently out-of-control signals may be expected to occur when a process is in control and why they will occur much more frequently when it isn't. The book introduces statistics from scratch, making no assumption that the reader has been exposed to fundamental statistical ideas or to statistical measures.

Chapter 5 is devoted to a deeper understanding of probability. The first section examines basic rules used for computing probabilities and conditional probabilities and shows how probability may be used to answer questions in business, industry, and other areas. The second section details some of the probability distributions used most often to answer questions of interest in the quality arena. Along with being able to apply these distributions in a variety of settings, the reader will gain further expertise with and a deeper understanding of control charts. The third section is devoted to the topic of reliability of a product or system. Frequently used reliability measures are introduced—for example, the reliability function, failure rate, and mean time to failure.

Using statistics and probability properly, while essential in making good process decisions, is not the whole of the quality picture. A company can optimize how it uses statistics and still fail because of its management process, how it treats customers or employees, or how it solves problems.

Chapter 6 is devoted to *thinking* quality. Topics include the management philosophies of Deming and the thinking of other quality visionaries such as Juran and Ishikawa. Additionally, there is an overview of total quality management and of the Six Sigma movement. Problem-solving tools are introduced—for example, Pareto charts, cause-and-effect diagrams, scatter diagrams, flowcharts, and methods a group can use when it can't reach consensus on choosing one option from among three or more.

This book grew out of two sources—a course I designed and taught while consulting with a manufacturing company, and a course I designed and taught for the MS Program in Mathematics at the College of Charleston, where I have "professed" mathematics and statistics for many years. This book has a definite "how to" point of view. I repeat. This is a practical book.

Portions of the input and output contained in this book are printed with the permission of Minitab Inc. MINITAB® and the MINITAB logo® are registered trademarks of Minitab Inc. The reader who has access to Excel or MINITAB will have some opportunities to use them, and even instructions for how to do so. However, no software is required to understand this book. Occasional exercises that are included for those who have such software may be safely skipped.

I wish to thank some good people. A number of years ago, Jim Lynch invited me to sit in on an SPC short course when I was on sabbatical at the University of South Carolina. Seeing how such a course could work has benefited my teaching in many ways. Several years later, I was able to spend a sabbatical year working on SPC issues with a manufacturing company. Solving mathematical and statistical problems that help a company make a product better gives a satisfaction that is hard to top and, well, is just plain fun. Gordon Jones, then-Dean of the School of Sciences and Mathematics at the College of Charleston, liked the idea of such a sabbatical and approved it. Way to go, Gordon. I have consulted with some wonderful people on projects for various companies, but particularly want to acknowledge Gregg Adams, Neal Tonks, and Ben Bruner. Also, thanks are due to Debbie Yarnell at Prentice Hall and the following reviewers for suggesting additional topics and ideas that would improve the scope of the book: J.K. Crain, Texas A&M University—Commerce; Grace Duffy, Trident Technical College; Mark Durivage, Owens Community College; Samuel H. Huang, University of Cincinnati; Young J. Kim, Mississippi State University; Kellie Knox, Southwest Wisconsin Technical College; Jooh Lee, Rowan University; George Pillinayagam, Lorain County Community College; and Carl Wargula, Gateway Community College.

My son Andrew didn't squawk much when I wanted him to quit running over pedestrians and running into other vehicles so that I could use the computer. My daughter Susan tended to come in the room whenever I happened to be typing a personal or family anecdote for a homework problem that asks the reader to identify a principle of TQM or one of Deming's

14 points that is being violated. I would ask her to read the problem and whether she remembered the outrageous incident. She would always read the problem, say she remembered, then invariably tell me that the purpose of her visit was to let me know that she was on her way out the door to do this or that. Thanks for your patience, Susan. We're glad you're you. Last but definitely not least is Libby, who has given me much encouragement toward working with industry and with this book. Additionally, when you're writing a book and the computer gets temperamental, like when you get the idea to paste all of the sections into one big file called *The Whole Enchilada* and the computer chokes, it is particularly helpful to have a wife who is a computer programmer. Really.

RMN

A Quick Course in Statistical Process Control

Contents

1

Statistical Preliminaries for Control Charts

OBJECTIVES

- Explain the difference between common cause and assignable cause variation
- Illustrate commonly made misinterpretations in statistics
- Define and show how to recognize overcontrol
- Explain fundamental statistical measures used in SPC
- Explain the Empirical Rule and when to apply it
- Explain the concept of random sample
- Explain how variation in a random sample relates to variation in the population that has been sampled

1.1 ADOPTING A STATISTICAL POINT OF VIEW

Statistical Process Control (SPC): *The use of statistical methods to control/improve a process.*

The general goal of SPC is to produce better goods and services. People are better off when the products they use last a long time and function reliably. Also, people who are customers—and that's everyone—are consistently happier when service providers know their stuff, treat them in a friendly, professional manner, and make them willing to be long-term customers. When data can be collected to measure the quality of a manufactured product or of a service, the potential is there to use the data to tell if quality is slipping or holding steady, and whether efforts to improve quality are working. The whole philosophy of SPC is to use data to continually improve quality. One of the twentieth-century giants of statistical process control, W. Edwards Deming, once said "In God we trust. All others must bring data." Of course, having data is not the same as understanding what the data reveal. This is where statistical methods come in. The basic statistical tools of SPC make a number of things possible.

One can *monitor a process to determine if it is working to the best of its capability.* Any system will degrade without proper oversight and care. Quality may be slipping or holding steady, and the tools of SPC, such as control charts, can determine which.

Generally, quality does not improve by accident. Improving the quality of manufactured goods or of services means making changes in the system. ***Control charts*** can be used to *determine if intentional changes in the process are having the desired effect.* For example, has the new drill press reduced the proportion of drilled holes that don't meet specifications? The fact that there may be fewer ***out-of-spec*** holes produced the day after the new drill press was put to use than the day before doesn't make a statistical case that the process has improved. *Too many people want to take credit because today's numbers are better than yesterday's, or worse, blame someone because today's numbers are not as good as yesterday's.* Even the same drill press would be expected to produce a different number of out-of-spec holes from day to day. The same customer service representative will not generate exactly the same satisfaction level from one customer to the next. Of course, it is also possible that today's number is so much better than yesterday's that the difference is beyond reasonable chance. But, how big of a difference is required to be sure that the positive effect is beyond reasonable change? It is also possible that the combined weight of several days-worth of numbers, no one of which is compelling by itself, can indicate improvement beyond reasonable chance. Control charts can signal when either level of improvement has occurred.

The previous point indicates that understanding variation in numbers is important. Control charts and summary statistical measures we will discuss, such as the range and standard deviation, can be used to *detect when there are unwanted sources of variation in a process.* No two items made are perfectly alike. In any production process there is always variation from item to item, batch to batch, at different points in a continuous flow, and so on. Some amount of variation is impossible to avoid. Still, in business and industry, variation is the enemy. Whether there is *needlessly high* variation from product unit to product unit is a different question than whether data will show that the average product unit is on target. A manufacturer wants to reduce unit-to-unit variation so that customers will come to appreciate the *consistency* of the product. Monitoring variation is important so that it may be kept to a minimum. Control charts and statistical measures are tools used in this effort.

Additionally, SPC has methods for measuring how well a process is performing—that is, measuring what is called the ***capability of the process.*** It is natural to ask how capable a process was this week, this month, during a particular product run, and so on. By examining the data from different time periods one can determine if the process is more or less capable, currently, than it used to be. An actual statistical measure of process capability is called a ***process capability index.*** It is a very common practice for Company A to ask to see Company B's capability indices for a particular time period. The purpose is to help Company A decide whether it should purchase goods made during that period by Company B. The corollary is that to be competitive, Company B needs to be collecting the kind of show-and-tell data that produce control charts and capability indices. Needless to say, companies that want access to the largest markets are companies that make continuous efforts to improve process capability.

Uses of SPC, such as those just discussed, are tied to two related objectives:

- continuous improvement in product and process, and
- increasing product and process consistency by reducing variation.

Many companies prominently display posters with slogans such as "Continuous Improvement," "Zero Defects," "Get It Right the First Time," or "Work Smarter, Not Harder." But these slogans are useless if the posters are seen but don't reflect how employees think or act. Positive things happen when continuous improvement is a workforce state of mind with action consequences. This equates *continuous improvement* with *never being satisfied with how things are.*

Merely having production meet product specifications should not be viewed as good enough. If the average product measurement is off target this week, what can be done to bring the process closer to target next week? What process changes can be made to reduce variation so that the product will be more consistent? Remember: Variation is the enemy. The less variation there is, the less rework there will be, the less off-spec production there will be, and the less scrap there will be. Can standard operating procedures be changed in order to speed up the change-out of parts? To reduce downtime? To reduce the frequency of laboratory tests? To reduce the incidence of mislabeled shipments? or to reduce the drain on workers caused by morale busters? Questions like these reflect the thinking process that accompanies the tools and practices of SPC. This state of mind has the best chance of prevailing in the workplace if employees see that management has acquired it first and "walks the talk." This type of thinking gets at the "soft," or philosophical, side of SPC, which is introduced in Chapter 6.

So where does statistics fit into the picture? The answer is that understanding and controlling variation is what statistics is all about. We will use tools such as the *standard deviation* and *range* to measure the amount of variation present in a process. This will enable us to set up control charts and then use them to determine if variation is stable or changing—for the better or the worse.

Variation can be viewed as having one of two causes:

> ***Common cause variation*** is variation inherent in the process. This variation is natural in a process that is working as it was designed to work.
>
> ***Assignable cause variation*** (aka *special cause variation*) is variation attributable to a source present that makes the process not work as designed.

Even when a process is functioning at its highest capability, there will be variation. The resistance in consecutively produced insulators will be different if measured to enough decimal places, two ball bearings will have slightly different diameters, two customers who order the same thing will wait in line different amounts of time at a fast food restaurant, a clerk will occasionally misplace an order or enter the wrong number on a customer order form, and so on. The variation in a system that is working up to its design capability is *common cause variation*. Some level of variation has to be tolerated in any system—not wanted, just tolerated. Further, there should be a never-ending effort to reduce this kind of variation. But one can never expect to completely eliminate variation. Even the best of people sometime make mistakes. No system is perfect.

One of the key purposes of control charts is to signal when *assignable cause variation* may be present. In order to set up a chart that can do this, one first must measure the level of common cause variation in the process. This brings us to the central idea of why control charts work:

> Knowing the level of common cause variation in a process makes it possible to define limits that, when exceeded, suggest that one should investigate whether an assignable cause is present.

How Not to Think—An Example

The following data are taken from an AP article that appeared in the Charleston, South Carolina *Post & Courier* in June of 1991 (see the article "Senate Panel Cites 'Plague' of Rural Criminal Violence"). A U.S. senator who will go unnamed proposed spending $100,000,000 to investigate a "plague" of crime in the rural United States. One of a number of faulty "statistical" justifications given in the article was the following comparison of the increase in crime in Montana, an example of a rural (that is, low population density) state, to Los Angeles, an example of a high population density region:

Murders	1989	1990	Increase
Montana	23	30	30%
Los Angeles	877	983	12%

The idea here is that when the murder count in Montana increased by 7 from one year to the next, that amounted to a 30% increase, whereas during the same time frame, the count in Los Angeles increased by 106, just a 12% increase. What the article attempts to convey is that a 30% increase is more alarming than a 12% increase, which supports the idea that a plague of crime exists in the rural United States.

There are several problems with this reasoning. One is that one year's number is being compared to the previous year's number. Depending on the numbers, a statistician might be able to make a compelling argument about a Montana crime problem if murder counts had risen systematically over a succession of years rather than just over two years. In the absence of more than two year's worth of data, a case still could be made with just two year's worth of data, as we shall see, but it would not be made on the basis of percent increase. The problem with using percent increase is that when working with small numbers, small changes can translate into big percent increases. An increase of 7 murders amounts to a considerable percentage increase because the reference point is 23 murders. By the article's reasoning, an even stronger case could be made for a rural state that had one murder in one year followed by two the next. The additional one murder represents a whopping 100% increase. In the meantime, the additional 106 murders in Los Angeles represents only a 12% increase because the reference point is 877. The issue is whether a murder increase of seven is unusual. To decide this, we must first understand how much fluctuation is typical from one year to the next. Once the amount of typical fluctuation is assessed, it will be possible to define a limit that, when an annual murder count exceeds

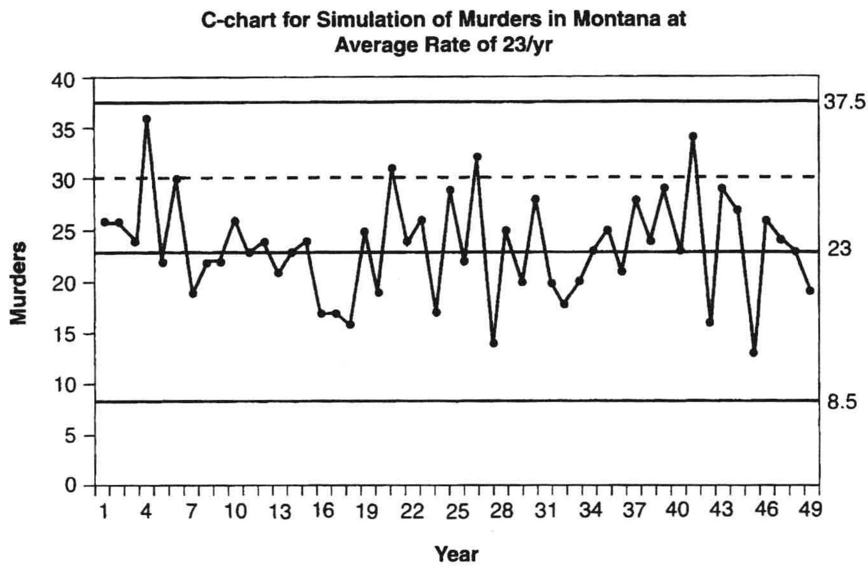


FIGURE 1.1 *C*-Chart for Simulation of Murders in Montana at Average Rate of 23 per Year.

it, suggests that one should investigate whether an assignable cause is present. We can illustrate this concept by looking at our first example of a control chart.

Figure 1.1 is an Excel chart that represents how 50 consecutive years of murder counts in Montana might graph under the assumption that the expected number of murders in any one year is 23. Many statistical computer packages can generate, that is, *simulate*, such data based on probability models. The numbers produced by these models, the number of events of a given kind that occur in each of a succession of time periods—in our case, the number of murders each year for 50 years—will closely mimic the kinds of numbers that nature produces. To generate such data, all one needs to know is the expected event rate per time period. In our case, this is the annual murder rate. Since no one was alarmed in the year when 23 murders occurred, we take the annual murder rate to be 23, and look at the simulated data to help determine if 30 or more murders in a year is unusual. The important thing here is that Figure 1.1 shows the kind of year-to-year variation expected from a system that is perfectly stable over the course of many years.

Figure 1.1 is an example of a *c*-chart, which we will meet in more detail later. At this point, we simply note that the chart has three horizontal reference lines. One has a *y*-intercept equal to the average annual murder rate, 23. As expected, some years have more than 23, some less, and there appears to be a random mix of points above and below this line.

The other two lines are used to identify any points that are unusually far from 23. The *y*-intercepts of these lines, 37.5 and 8.5, are called ***control limits.***

Later we will see how these limits are determined. For now it suffices simply to understand how the limits might be used in a manufacturing process.

The *c*-chart is commonly used to monitor, over a sequence of time periods, the number of ***nonconforming*** items produced, or items which do not ***conform*** to specifications. Some examples are:

- the number of defective bumpers produced in a day,
- the combined number of blemishes in the paint jobs of a three-car sample (the sample should be considered as representative of all the cars painted on a particular day),
- the number of accidents in a month, and
- the number of flaws in each 1,000-yard sample of copper wire insulation (the 1,000-yard sample should be considered as representative of all the wire produced on a particular 12-hour shift).

When a point occurs that does not fall between the control limits, the event is considered so unusual that the manufacturer would investigate to determine if there is an assignable cause that can explain the unusual count. If an assignable cause is found, any problem must be fixed so as not to affect future counts. If the problem is serious enough, the manufacturing process may need to be shut down, process permitting. Shutdowns are never good for the company, its profits, or its employees.

On most control charts the control limits are determined so that, by chance, only about 3 points in 1,000 would be expected to not fall between the control limits when the process is working up to its capability. As indicated, a process that is working as it was designed to will sometimes produce points falling outside the control limits. But this is rare, occurring roughly three times in one thousand, on average. If no assignable cause can be connected to such a point, it is assumed that the point is one of those "three in one thousand," and the process is allowed to continue.

Note, too, that there is a probability symmetry for the control limits. When the process is working up to its capability, only about 1.5 points in 1,000 would be expected to fall *above* the larger of these (called the ***upper control limit [UCL]***), and about 1.5 would be expected to fall *below* the other (the ***lower control limit [LCL]***).

We now return to Figure 1.1 and the question of whether 30 or more murders in a year is unusual when the expected murder rate remains a constant 23 over a span of years. A reference line, a dashed line with *y*-intercept of 30, is shown in the figure. The murder count is 30 in year 6, and exceeds 30 in four other years. There are also three other points "knocking at the door," with 29 murders in those years. In any case, in one-tenth of the years, the count is 30 or more. In statistics, an event that occurs in one-tenth of repeated experiments is not considered to be unusual. Of course, had we simulated another 50 years of murder counts, the fraction of years in which there are 30 or more